



This is a repository copy of *Intelligent Genetic Algorithms in Evolutionary Computation Part 1. Biological Foundation*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/83548/>

Monograph:

Liwen, He. and Mort, N. (1998) *Intelligent Genetic Algorithms in Evolutionary Computation Part 1. Biological Foundation*. Research Report. ACSE Research Report 689 . Department of Automatic Control and Systems Engineering

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

x

Intelligent Genetic Algorithms in Evolutionary Computation

Part I. Biological Foundation

Liwen He and Neil Mort
Department of Automatic Control and Systems Engineering
University of Sheffield,
Sheffield, S1 3JD, UK

Research Report No. 689

Abstract: In this paper, we review a large amount of historical biological literature, [Darwin, 1862, 1871; Fisher, 1930 and others] and recent developments in biological [Anderson, 1994] and biocomputational literature [Miller & Todd, 1992, 1994], try to integrate the dynamics of interplay between natural selection and sexual selection through mate choice in biology with evolutionary computation as a process of search, diversification and optimization, and originate a new class of evolutionary algorithm which we term Intelligent Genetic Algorithms. These intelligent genetic algorithms demonstrate their effectiveness and efficiency in generating evolutionary innovations, maintaining genetic diversity, promoting mate choice and sexual recombination in species, and guiding the movement of a population from local optima to global optima in parallel. Furthermore, we attempt to provide some common biological origins for these new Intelligent Genetic Algorithms.

Keywords: Evolutionary Computation, Genetic Algorithms, Fitness Landscape, Natural Selection, Sexual Selection, Phenotype and Genotype

Contact Details (both authors):

☎ +44 114 222 5601

FAX +44 114 222 5602

✉ n.mort@sheffield.ac.uk

April 1998

200425453



1. Introduction

The most complex and elaborated life-forms on the earth are flowering plants and animals with neural systems. They share the common characteristics that both carry out sexual selection through mate choice. Flowering plants are sexually selected by heterospecific pollinators such as insects and hummingbirds that they attract to further their own reproduction (Darwin, 1862; Barth, 1991). and animals are sexually selected for reproduction by opposite-sex conspecifics (Darwin, 1871; Andersson , 1994).

Darwin (1862) suggested that the evolution of phenotypic complexity and diversity may be driven not simply by natural-selective adaptation to niches, but by a complementary interplay of natural selection and sexual selection . Though natural selection is the greatest power to propagate evolutionary success, sexual selection is a powerful mechanism to explore novelty, diversification and optimization .

Natural selection normally leads to evolutionary convergence to unimodal (global or local) optimum because organisms adapt to the ecological environment automatically. So the natural selection method is a wonderful hill-climbing search method and good at exploiting adaptive fitness peaks.

On the other hand, sexual selection often leads to an unpredictable, divergent evolutionary procedure, with lineage speciating spontaneously and exploring phenotypic regions under the guidance of their capariously evolved mate choice. The complex and chaotic dynamics of sexual selection results from the fact that a species evolving through strong selective mate choice is a self-organizing system that can determine its own evolutionary trajectory [Miller, 1992].

Therefore it is easily understood that the interplay between the dynamics of natural selection and the dynamics of sexual selection through mate choice can help promote both the exploitation of adaptive peaks and the exploration of novel regions of the search spaces, which is a central problem in GAs.

This interplay can be viewed as a critical factor that helps populations create new evolutionary innovations, preserve evolutionary diversification, generate automatic speciation, and search for new adaptive spaces in parallel, thus resulting in faster convergence to global optimization in evolutionary computation algorithms.

Some definitions that will be used in this paper are given in the sequel:

DEFINITION 1. Let $l \in \mathbb{N}^*$, be a positive integer, and let $\Omega = A^l$, with $A = \{0, 1\}$, be a *Genotype Space*. A *chromosome* is an element S of Ω , and l is the length of chromosome. *Gene* is the position of chromosome. *Allele* is a different form of gene .

The term $F: \Omega \rightarrow \mathbb{R}_+^*$, (\mathbb{R}_+^* is the set of positive real numbers) is called a fitness function map. In genetic algorithms, $F(S)$ is the fitness value of $S \in \Omega$.

DEFINITION 2. A chromosome with a genotype g_m is an element of genotype space Ω (l is the dimension of Ω) with

$$g_m = (a_1^m, \dots, a_l^m)$$

where a_i^m is the *allele* associated with the i th gene of chromosome m .

The *Hamming Distance* (Feistel, R. et al, 1989) between two chromosome g_m and g_n is:

$$d(g_m, g_n) = \min_j \sum_{i=1}^{l-j} \delta^*(a_i^m, a_{i+j}^n)$$

(with δ^* being the dual Kronecker function : $\delta^*(a, b) = 1$ only if $a \neq b$). This distance represents the minimal number of non-coincidences between the two chromosomes. For example, if $g_m = (1\ 0\ 0\ 1\ 0)$, $g_n = (0\ 1\ 0\ 0\ 1)$, then their Hamming distance $d(g_m, g_n)$ is 4.

DEFINITION 3. With the Hamming distance as a metric, Ω can be represented as a multi-dimensional lattice on which the mutation of one allele defines immediate neighbour (*1-mutant*) or the mutation of two alleles defines intermediate neighbour (*2-mutant*). So a chromosome has $\binom{l}{1}$ 1-mutant and $\binom{l}{2}$ 2-mutant.

DEFINITION 4. Since each genotype S can be assigned a real valued fitness $F(S)$, a *fitness landscape* can be pictured as an $(l+1)$ - dimensional plot in each genotype is a point in l dimensions and its fitness is plotted the $(l+1)$ st axis. (Mitchell, 1996)

Epecially, if each genotype S can be assigned a vector $Fv(S)$, we name it as *relative fitness vector*, whose dimension is p , a *relative fitness landscape* can be pictured as an $(l+p)$ dimensional plot in each genotype.

DEFINITION 5. Let $B = \{0, 1, \#\}$ where $\#$ is a wild card and or "don't care" symbol, it represents either value 0 or value 1. A *genotype schema* is an element of B^l .

For example, if $l = 3$, then $\{010, \#11, 0\#1, \#\#\#\}$ are schemata of length 3.

DEFINITION 6. A mapping transforms the *genotypes* into *phenotypes* which describes the physical properties of the associated individuals. For example, each phenotype can be described by a set of properties such as an individuals' size, weight, color or construction.

DEFINITION 7. Pareto Optimality (Goldberg, 1989): we state that a vector x is partially less than y , symbolically $x <_p y$, when the following conditions hold:

$$x <_p y \Leftrightarrow (\forall_i) (x_i \leq y_i) \wedge (\exists i) (x_i < y_i)$$

under these circumstances we say that vector x *dominates* vector y . If a vector is not dominated by any other, we define it as *nondominated*.

Specifically, if there are three individuals A, B, C in a population which is characterized by the binary vector $Va=[0\ 1\ 0\ 1\ 1\ 0\ 1]$, $Vb=[0\ 1\ 0\ 0\ 1\ 0\ 1]$, and $Vc=[1\ 1\ 0\ 0\ 1\ 0\ 1]$; it is easy to see that individual B is dominated by both individual A and C, individual A and C are not dominated with each other. If individual A or C are not dominated by any other individuals in the whole population, they are *nondominated individual*.

Also, the following characteristics of IGAs are developed throughout this paper:

- 1) *Spatially Structured Population* searches for multimodal optima in the fitness landscape.
- 2) Individual's *relative fitness vector* based on the population's phenotype information helps identify non-dominated individuals.
- 3) *Super Conservative Selection Strategy* helps preserve non-dominated individuals in the next generation, thus maintain genetic diversity.

4) *Dynamic Sharing Fitness Function* favors novel and distinct individuals, then prevents population from premature convergence.

5) '*Hybrid drift*' model can help explain the mystery how populations can achieve success in escaping from local optima to global optima.

6) *Intelligent Recombination* and *neutral mutation* can produce more viable, fertile and productive offspring, and increase the speed and robustness of finding global optima.

2. What is the effect of 'mate choice' in evolution ?

As we know, the two basic functions of the animal nervous system are: (1) generating adaptive survival behavior that finds food and water and avoids predators , and (2) generating adaptive reproductive behavior that exploits the availability of viable, fertile and attractive mates .

Mate choice has two forms: *assortative mating* and *selective mating* (Miller, 1992). *Assortative mating* is a mechanism where common species are clustered together, recombination can only be carried out within similar organisms, and there is no outbreeding across species, thus reinforcing the speciation and parallel search, and enhancing the offspring's exploitation around their fitness peak. On the other hand, selective mating is a way of mating with individuals with a higher fitness value, allowing offspring to obtain the most favorable, viable and productive traits from their parents, which in turn increases an offspring's fitness value.

Thus mate choice is a way of "internalizing" (Miller, 1992) historical selection pressure into the current phenotype search space, and guiding the offspring along an evolutionary trajectory in a more 'intelligent' directional movement towards adaptive peaks. Mate choice is a central evolved mechanism which we can learn from nature and implement in evolutionary computation. For example, assortative mating enlightens us on how to design a population structure and build up sub-populations, while selective mating can help us choose or design a proper selection strategy and appropriate genetic operators.

3. Evolutionary Innovations

Evolutionary innovations are distinct characteristics that distinguish one species from another. A complete list of evolutionary innovations might include at least all diagnostic features of all species in the earth. Classical examples of major evolutionary innovations include the bony skeleton of vertebrates, the jaws of gnathostomes, the amniote egg, feathers, continuously growing incisors, large brains in hominids, and so on. (Cracraft, 1990). Evolutionary innovations are important because natural selection creates adaptations out of innovations: "Innovation is the mainspring of evolution" (Jablonski & Bottjer, 1990, p253).

However, the question of "what is the power driving evolutionary innovations?" has been under discussion for more than a century. Fisher (1930) followed Darwin (1883), and indicated that innovations could indeed be produced purely through natural-selective hill-climbing. The main problem with this idea is that it ignores the condition of local optima; however, and Goldschmidt (1940) considered that macromutations

could generate fully functioning novelties in the form of "hopeful monsters". As we know, 'mutational search' is a random, passive, unpredictable way to direct population escape away from local optima. Mutation might not be the main contribution of innovations in nature. Miller and Todd (1992, 1994) suggested that sexual selection through mate choice can account for all of these features of evolutionary innovations, where sexual selection is to macroevolution what genetic mutation is to microevolution, significant innovations may originate from large amount of parallel searching through a multi-dimensional adaptive fitness landscape.

Based on these previous viewpoints, we suspect that evolutionary innovation should be explained by the macroevolution of interplay of natural selection and sexual selection through mate choice with mutation. When populations stick to local optima, divergent sexually-selective dynamics prioritize in exploring a prospective zone of the search space; if such a zone is encountered, naturally-selective dynamics are responsible for the directional search and hill-climbing, an evolutionary innovation is generated, mutation makes its own contribution to the 'adaptive walk' new area surrounding the fitness peak under the guidance of interaction between natural selection and sexual selection.

4. Diversification through Automatic Niching and Speciation

Mayr (1942, p120) defined species as "groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups". Eldredge (1989, p99) advised that diversification represents the gradual accumulation of various phenotypic adaptations. In the model of *allopatric* speciation, Mayr (1942) suggested that speciation can be created by geographic barriers. In the model of *sympatric* speciation, Dobzhansky (1937) suggested a more abstract cleaver: a low-fitness valley in an adaptive landscape. Both of these models tend to cleave populations apart with a causal and deterministic force external to the speciating populations themselves. In the Quicksilver Model of speciation, Todd and Miller (1991) showed that sympatric speciation can occur in the absence of external divisive force; mate choice can spontaneously differentiate in a population through a combination of genetic drift, mutation pressure, and assortative mating, such that speciation can occur even without geographic isolation or ecological specialization.

In nature, large numbers of species can exist because species diverge ecologically and/or geographically. Therefore evolutionary diversification can be produced by combining spatially structured population, ecological niching, and assortative mating, then letting populations decide by themselves how to speciate and construct spatial sub-population automatically.

The spatially structured population can facilitate diversification in evolutionary computation. In intelligent genetic algorithms we construct a kind of spatially structured population based on the similar phenotypic schemata: around any schema, there is a sub-population whose individuals all include this schema. So the number of sub-populations is decided by the number of phenotypic schemata discovered so far. Unlike the spatially structured populations of island model or fine-grained model in parallel genetic algorithm (Dorigo.M & Maniezzo.V, 1993), the sub-population in intelligent genetic algorithms is semi-isolated, that is to say: any individual with more

than one phenotypic schema will be the member of more than one sub-population, the more phenotypic schemata it has, the higher fitness value it has, and the more subpopulations it belongs to. The cleaver of sub-population is naturally and automatically formed. Evolved populations 'know' how to split into sub-populations by themselves. This spatial population structure should be reliable and efficient to produce genetic diversity, explore multiple optima in a high dimensional landscape for most combinatorial optimization problems. It is clear that this spatially structure population can be easily implemented in parallel computers for large problems.

Another problem is how to maintain speciation. In the earth, different species do not go head to head, but live a coexisting life on their own ecological nich.

In intelligent genetic algorithms, we develop a 'super conservative selection strategy', where different non-dominated individuals are kept in the next generation after selection until they are dominated by other species with higher performance. Unfortunately, traditional GA theory has placed the burden of preservation as well as vigorous recombination completely upon crossover, forcing a trade-off between preservation and exploration. Operators that tend to preserve schemata also tend to lead to rapid convergence because of small population size [Eshelman & Schaffer, 1991]. The usual method to keep genetic diversity and slow down premature convergence is the 'Sharing' method, but this makes crossover more disruptive of good schemata, thus disrupting the effectiveness of processing schemata in parallel. Fortunately, intelligent genetic algorithms use 'super conservative selection strategy' to keep nondominated individuals in the next generation, just to keep the distinct schemata searched. This strategy transfers the burden of maintaining useful schemata from crossover to selection; so the 'highly disruptive' crossover operator can be used to explore a broader search space and find prospective zones which may lead to a new fitness peak.

In a previous study [He & Mort 1997], when a "combination" for an individual is searched, the fitness value of the individual is increased by one, thus the fitness space is very 'rugged'. If an individual is honored with a higher fitness value than other individuals randomly, its offspring will increase exponentially, and be dominant in the whole population. Useful schemata are lost in the early stages because of small population size, and genetic diversity is limited. This results in premature convergence.

In the continuous multimodal optimization problem [Goldberg and Richardson, 1987; Deb and Goldberg, 1989], a 'sharing method' is used to induce speciation for the individuals by measuring their physical distance from each other and determining their similarity. This method needs a priori knowledge of the fitness landscape in order to choose suitable parameters.

In intelligent genetic algorithms, we initialize the 'dynamic sharing fitness function' to favor the novel schemata, preserve genetic diversity, and propagate the speciation automatically. The 'dynamic sharing fitness function' (fis) is stated as the product:

$$fis = (fa) (\gamma^{sh})$$

where fa is the physical fitness value of individual, γ is a constant rate term, and $0 < \gamma < 1$; sh is the similarity coefficient of an individual relative to all other individuals in the entire population. The dynamic sharing fitness function indicates that the real fitness value of an individual is decreased with the increase of its similarity with other individuals.

This method is especially suitable for solving the combinatorial multimodal optimization problem where very little information about the fitness landscape is known *a priori*.

The 'dynamic sharing fitness function' favors the new schemata searched and biases the distinct individuals, prevents the individuals with higher fitness value being dominant in the population, tries to put individuals into separate niches like the ecological 'exclusion principle' (two species cannot occupy the same niche in the same area and at the same time), and distribute individuals across fitness landscape in proportion to the height of their peaks. Reproductive isolation through speciation can consolidate the adaptive differences between individuals and lead to ever-increasing biodiversity and niche differentiation [Todd & Miller 1991].

5. Global Optimization through Natural Selection and Sexual Selection

Given that natural selection is a good hill-climbing search mechanism, why does nature use selective mating as a complementary search tool for optimization ?

First, let us consider ecological local optimization through fitness-based sexual selection.

When operating in the wild, natural selection is often a noisy, irregular, and inaccurate process [Miller, 1992]. For example, predators might eat the prey that has the better vision, larger brain, and longer legs, simply because that animal was more easily noticed or more active at feeding time than the duller, blinder, slower animal. Natural selection itself has no incentive to reduce this noise. On the other hand, selective mate choice can provide enough information about the phenotype quality of potential mates. As Darwin (1871) stated, females are usually choosier than males about their mates. Female animals convert slight differences in male quality into large differences in male reproductive success by mating with males with local optimal fitness value. Males become the explorers of the adaptive landscape, they are forced by the female to search through a larger phenotypic space in order to satisfy the female with sexy, colorful, novel and attractive courtship behavior. Once the males happen to encounter a new adaptive zone close to a higher fitness peak, a combination of natural selection and sexual selection may drive the entire populations, including males, females and their offspring, to climb new higher peak and generate new evolutionary innovations. This more accurate local optimal information helps offspring reduce noise and evolve precisely on their own evolutionary trajectory. Thus selective mate choice is an automatic and natural fitness scaler that matches the fitness landscape intelligently [Miller, 1992], and tends to impose a kind of 'directional preference' to move populations along an evolutionary trajectory in a particular direction to global optima.

Second, populations easily stick to local optima in evolutionary computation. This naturally begs the question "How can populations succeed in escaping away from local optima to global optima in nature?"

Wright (1932, 1982) in his "shifting balance" theory suggested that *genetic drift* running in quasi-isolated populations can sometimes enable one population to move far enough away from its current fitness peak and to enter a new prospective area around a higher fitness peak. This increases evolutionary diversification and brings about new evolutionary innovations when climbing this new peak. Our previous experience in

simple genetic algorithms (He & Mort, 1997) demonstrated that in a very rugged landscape, "adaptive walks" with short steps (defined relative to the landscape's ruggedness) generated by genetic drift and mutations, allow individuals to move from local optima. In that scheduling optimization problem, when the mutation rate is chosen to permit only one bit flip in the binary, a single optimum is obtained. So we propose that genetic drift and mutation may be the main forces that drive population escape from local optima to global optima in simple genetic algorithms .

Miller and Todd (1992) suggested a kind of 'sexual-selective drift' from the stochastic dynamics of mate choice and runaway sexual selection (Fisher, 1930) to drive populations away from local optima, and suspected that this sexual-selective drift would be stronger and more directional than simple genetic drift.

Another interesting theory is Kimura's (1983) 'neutral drift' through adaptively neutral mutations, which is usually applied to molecular evolution (DNA base pair substitutions typically do not change protein functionality, but the gene construction of protein is changed in reality). Under this model, populations can drift around many equal-fitness-value 'ridges' and explore new areas through the high-dimensional adaptive landscape without incurring fitness cost. Assuming this condition, natural selection has no effect on this process, while neutral drift lead populations from one adaptive zone to another, expecting to find a new adaptive zone associated with higher fitness peaks.

However, we believe that in such a complex world, the difficulty of sticking to local optima is not easily solved by any one kind of simple model. We are trying to build up a '*hybrid drift*' model to overcome this problem. Consider the evolution of monkey to mankind as an example. At first, the transfer from quadrupedalism to bipedalism for monkey is just a specific, accidental, novel behaviour, also reinforcing the evolutionary diversification in the entire monkey population. From the viewpoint of fitness landscape, quadrupedalism and bipedalism are regarded as two equal-class fitness peaks. No one knows which is better beforehand, so no natural selection operates on them. Through sexual selection with mate choice, females favor the mates displaying novel courtship behavior and happen to mate with a bipedal male, then female, male and their offspring, form a special sub-population, and all enter a new adaptive fitness zone through this kind of 'neutral drift'. At this point, natural selection could act on this sub-population: Over a long period of evolution, this 'standing' sub-population can observe a richer world, facilitate their upper limbs to use tools to obtain food, and manage to adapt to the environment intelligently. This promotes the development of their brains, and evolutionary innovation is generated. Genetic drift tends to spread this bipedalism with higher fitness value in the entire population of relatives. In the end, mankind with the most complicated 'human brain' and highest intelligence in this planet appears. Therefore, the 'hybrid drift' model through the complementary interaction between natural selection and sexual selection may explain the mystery that populations can be internalized to escape away from local optima (monkey) to global optima (human being).

In intelligent genetic algorithms, we utilize the 'hybrid drift' model to design the populations' *relative fitness* value as a multi-dimensional vector according to their phenotypic information. When the relative fitness vector of one individual is non-dominated by that of any other one in the entire population, 'neutral drift' is in effect,

this individual is ranked as first class which means that it is guaranteed to have at least one offspring in the next generation through 'super-conservative selection'. In other words, any non-dominated individuals are regarded as having equal-fitness value, 'neutral drift' can enable distinct individuals explore different adaptive peaks around themselves. When one individual is dominated by another, 'genetic drift' is in effect, this individual will die out and replaced by new non-dominated one. This model can very well explain how to lead populations to escape from local optima to global optima in intelligent genetic algorithms.

6. Intelligent Recombination through Optimal Outbreeding

Simple genetic algorithms recombine individuals randomly. Intelligent genetic algorithms require an additional selection process before recombination.

Todd & Miller (1991) developed probability-of-mating (POM) functions defined across an entire n-dimensional phenotype space representing mate preference.

In intelligent genetic algorithms, we craft a 'male tournament' scheme undertaken before recombination, where an individual is randomly selected as the 'female' from sub-population; then a number of individuals are randomly chosen as possible mates. The Mate Preference between the female and every candidate is calculated according to the 'relation function'. Then the candidate with highest 'Mate Preference' value is the winner of this tournament, and is chosen as the 'male' to crossover with the 'female'.

The 'mating preference' function is defined as the product of phenotypic similarity (the number of similar phenotype schemata) and the genetic Hamming distance between two potential parents. When the binary representation of two parents is the same (the phenotype schemata is also the same) or their phenotype schemata is totally different, the mating preference function equals zero. This means that they cannot crossover.

This 'female chooses male' scheme incorporates more accurate direct phenotype and genotype information into evolutionary computation, allowing different phenotypically-separate and reproductively-isolated groups to preserve their own frequency distribution. This enhances speciation, and guides evolution more 'intelligently' towards optima based on previous knowledge.

In 1983, Bateson proposed the concept of 'optimal outbreeding', where animals prefer the somewhat similar to the exactly similar. An inverse example is that: in the Middle Ages, in order to keep their lineage pure or for political reasons, the imperial families in Europe always preferred marriage within the imperial lineage, i.e. marriage was limited to between cousins. These close relationship marriages resulted in their offspring having very low intelligence and a variety of hereditary diseases, thus making the imperial family die out. In nature, 'optimal outbreeding' overcomes this 'incest' problem, and tends to preserve good properties from parents, and avoids the danger of accepting the common poor features from two parents.

In evolutionary computation, recombination through optimal outbreeding is utilized to maintain useful schemata and building blocks, and also prevent premature convergence.

Eshelman & Schaffer (1991) designed an 'incest avoidance' method where binary individuals can mate only if their Hamming distance is above a certain threshold. They

reported that 'avoiding incest' significantly increases the speed and robustness of finding global optima on 11 of 13 test functions.

Eshelman (1991) developed a highly disruptive uniform crossover operator, HUX, that crosses exactly half the non-matching alleles (what Booker (1987) calls the reduced surrogate), where the bits to be exchanged are chosen at random without replacement. HUX guarantees that the offspring are always the maximum Hamming distance away from their two parents. The flip side of HUX's disruptiveness is that it maximizes the chance of two good schemata, one from each parent, being combined in a offspring since half the material from each parent is chosen. Also, Eshelman used a non-traditional genetic algorithm, CHC, which combines a 'conservative selection' strategy that always preserves the best individuals found so far with this highly disruptive recombination operator HUX to produce offspring maximally different from both parents.

In intelligent genetic algorithms, we use the 'super conservative selection' strategy which always preserves the existing non-dominated individuals and the 'intelligent recombination' strategy that employs the highly disruptive crossover operator HUX. Since the crossover is only within sub-populations, and the 'mate preference' function between two potential parent is calculated beforehand, the recombination strategy is a suitable method that allows individuals to avoid breeding with others that are too similar, as this may generate useless offspring, and to avoid breeding with others that are too different which may generate unpredictable and risky offspring.

7. Neutral Mutation Combined with Intelligent Recombination

In nature, aside from preferring similarity and quality in mates, many female animals prefer novel, rare males. Darwin (1871) regarded neophilia (preference for novelty) seriously as an element in sexual selection, that could often lead to evolutionary innovations unrelated to ecological demands. Neophilia can be implemented in evolutionary computation by allowing 'female' individuals to sample several 'male' individuals and to favor the one that is most different.

Harvey (1992) implied that one core problem in evolution is how to move a population already converged to a locally optimal fitness peak somewhere else that may be better.

In intelligent genetic algorithms, every sub-population is mutated using a 'neutral mutation' operator before recombination. This 'adaptive walk' around the surrounding regions of the search space combined with the intelligent recombination method, brings about the most important innovations in the relatively focused sub-populations whose members are sharing so much similarity. They also help these sub-populations prevent premature convergence, run away from local optima, and explore new adaptive zones.

8. Parallel Search through Speciation

At the level of population, indicated in section 4, different animals pair up on the basis of their similarity to let lineage specialize in exploiting (in parallel) different adaptive peaks in the extensive and complex phenotypic space, without wasting effort on interbreeding across different species of animals.

At the level of the single individual, indicated in section 5, females are usually choosier than males about their mates. Males become the explorers of the adaptive landscape, they are individually required by the females to search through the phenotypic space around their relevant adaptive peaks in parallel in order to satisfy the female with attractive courtship behavior.

In intelligent genetic algorithms, spatially structured population is utilized to fulfill a parallel search at the population level (macroparallelism), and create a set of sub-populations doing parallel searches in the fitness landscape.; and the 'male tournament' scheme is utilized to fulfill a parallel search at the individual level (microparallelism). It is well known that parallel computation is faster, more robust, and efficient than serial computation especially when implemented in a parallel computer. The strong sense of parallel search in intelligent genetic algorithms plays a critical role in solving multimodal optimization problems.

Furthermore, parallel search at both the individual level and population level is limited in phenotype, and carried out through 'explicit parallelism' in genotype space: Intelligent genetic algorithms can process many building blocks or schemata more directly and explicitly. This form of 'explicit parallelism': is different from 'implicit parallelism' in simple genetic algorithms [Goldberg, 1989]. Genetic algorithms operate on the population of strings. In evaluating a population of n chromosomes, the genetic algorithm is implicitly estimating the average fitness value of the all schema that are present in the whole population, and increasing or decreasing their representation or offspring according to the Schema Theorem. This simultaneous implicit evaluation of a large number of schemata in a population is known as *implicit parallelism*.

The concept 'implicit parallelism' means that the fitness value of the whole population is calculated in order to obtain the average fitness value in the selection scheme, Genetic Algorithms can implicitly process all individuals in parallel, thus can process a large number of schemata while processing a relatively small number of strings..

9. Characteristics of Intelligent Genetic Algorithms and Future Prospects

On the basis of previous discussions, we can summarize the characteristics of Intelligent Genetic Algorithms:

- 1) Spatially Structured Population is naturally generated to form sub-populations, it can help speed up evolutionary optimization and reinforce the robustness of intelligent genetic algorithms' convergence to multiple global optima. This spatially structured population can be readily implemented in parallel computers by assigning different sub-populations to different co-processors for large problems.
- 2) Super Conservative Selection Strategy helps preserve genetic diversity, and enhance speciation and parallel search in evolutionary computation, thus accelerating optimization and increasing the robustness of evolutionary adaptation.
- 3) Constructing the individuals' relative fitness value as a vector facilitates a high-dimensional relative fitness landscape, promotes 'hybrid drift' when populations stick to local optima, and helps the population search new adaptive zones.
- 4) Dynamic sharing fitness function can prevent premature convergence, and transform the simple and 'rugged' physical fitness search space into a complex and 'smooth'

space where different peaks can be searched by different species in parallel. It can be regarded as a way of incorporating information about past and current fitness spaces into the evolutionary search procedure.

5) Intelligent Recombination through neutral mutation is made up by defining 'Mate Preference' to all potential mate pairs, choosing the pairs with highest 'mate preference' using a 'male tournament' strategy, neutral mutating them with suitable mutation rate, and crossing them over using the HUX recombination operator. This intelligent recombination method can produce more viable, fertile and productive offspring than general genetic operators found in simple genetic algorithms, and increase the speed and robustness of finding global optima. Also, the male tournament scheme is not limited to implementation in genetic algorithms, but also could be implemented in genetic programming, evolutionary strategy, classifier systems and artificial life systems.

6) 'Hybrid drift' model through the complementary interaction between natural selection and sexual selection may explain the mystery that populations can be internalized to escape from local optima to global optima.

7) Intelligent genetic algorithms can evolve along the evolutionary trajectory without prior task knowledge of fitness landscape, and also require only a few parameters setting.

The work which follows in Part II develops Intelligent Genetic Algorithms applied to a typical combinatorial multimodal optimization problem.

References

- Andersson, M. (1994): *Sexual Selection*. Princeton University Press.
- Barth, F.G. (1991): *Insects and Flowers: The Biology of a Partnership*. Princeton: Princeton University Press.
- Bateson, P. (Ed.), (1983): *Mate choice*. Cambridge, UK: Cambridge University Press.
- Booker, L. (1987): Improving search in Genetic Algorithms. *Genetic Algorithms and Simulated Annealing*, L.Davis (ed.), PP 61-73. San Mateo, CA: Morgan Kaufmman.
- Bourgine, P. and Snyers, D (1995): Lotka Volterra Coevolution at the Edge of Chaos. Artificial Evolution, European Conference. Lecture Notes in Computer Science 1063.
- Cracraft, J. (1990): The origin of evolutionary novelties: Pattern and process at different hierarchical levels. In M. Nitecki (Ed.), *Evolutionary innovations*, PP.21-44. Chicago: University Chicago Press.
- Darwin, C. (1962): *On the various contrivances by which orchids are fertilized by insects*. John Murray.
- Darwin, C (1871): *The descent of man, and selection in relation to sex (2 vols.)* John Murray.
- Darwin, C. (1883): *On the origin of species* (6th ed.). New York: D. Appleton & Co.
- Deb, K. and Goldberg, D. E. (1989): An Investigation of Niche and Species Formation in Genetic Function Optimization. *Proceedings of the Third International Conference on Genetic Algorithms*.

- Dobzhansky, T. (1937) . *Genetics and the origin of species*. Columbia U. Press. (Reprinted 1982).
- Dorigo.M & Maniezzo.V (1993), Parallel genetic algorithms: Introduction and overview of current research, *Parallel Genetic Algorithms: Theory and Application*, J.Stender (Ed.), IOS Press, PP.5-42.
- Eldredge, N. (1989): *Macroevolutionary dynamics: Species, niche, and adaptive peaks*. New York: McGraw-Hill.
- Eshelman, L.J., & Schaffer, J. D. (1991): Preventing Premature Convergence in Genetic Algorithms by Preventing Incest.. *Proceedings of the Fourth International Conference on Genetic Algorithms*. University of California, San Diego.
- Eshelman. L.J (1991): The CHC adaptive search algorithms : How to have safe search when engaging in non-traditional genetic recombination. *Foundations of Genetic Algorithms*. Gregory J.E.Rawlins (Ed.), Morgan Kaufmann Publishers, PP.265-283.
- Feistel.R and Ebeling.W, *Evolution of Complex Systems: Selforganization, Entropy and Development*. Kluwer Academic, Dordrecht, 1989
- Fisher, R.A. (1930): *The genetical theory of natural selection*. Oxford: Clarendon Press.
- Forrest, S. (Ed.). (1993): *Proceedings of the Fifth International Conference on Genetic Algorithms*, Morgan Kaufmann.
- Goldberg. D.E. and Richardson. J., (1987): Genetic algorithms with sharing for multimodal function optimization. *Proceedings of the Second International Conference on Genetic Algorithms*. Morgan Kaufmann.
- Goldberg. D.E. (1989): *Genetic Algorithms in Search Optimization, and Machine Learning* , Reading , MA: Addison Wesley.
- Goldschmidt, R.B. (1940): *The material basis of evolution*. New Haven, CT: Yale University Press.
- Harvey , I., Husbands, P., & Cliff, D.(1992): Issues in evolutionary robotics. In J.-A. Meyer, H.L.Roitblat, S.W.Wilson (Ed.), *Form Animals to Animats 2: Proceeding of the Second International Conference on Simulation of Adaptive Behavior* (pp364-373). Cambridge, A: MIT Press / Bradford Books.
- He Liwen and Mort Neil. (1997): A Genetic Algorithm Approach to the Optimization of Scheduling of Jobs on a Multi-Product Batch Processing Machine. IFAC Workshop on *Intelligent Manufacturing Systems*, Seoul, Korea. PP165-170.
- Jablonski, D., & Bottjer, D. J. (1990): The ecology of evolutionary innovation: The fossil record. In M.Nitecki (Ed.), *Evolutionary innovations* (PP.171-190). Chicago: University of Chicago Press.
- Kimura, M. (1983): The neutral theory of molecular evolution. In M. Nei & R.K. Koehn (Ed.), *Evolution of genes and proteins*, PP.213-233. Sunderland, MA: Sinauer.
- Mayr, E. (1942): *Systematics and the origin of species*. Columbia U. Press. (Reprinted 1982).
- Miller, G.F. & Todd, P.M (1992): The Role of Mate Choice in Biocomputation: Sexual Selection as a Process of Search, Optimization, and Diversification. *Proceedings of The*



- Monterey Biocomputation workshop*, Monterey, CA. *Lecture Notes in Computer Science*, 899.
- Miller, G.F. (1994): Exploiting Mate Choice in Evolutionary Computation : Sexual Selection as a Process of Search , Optimization , and Diversification. *Evolutionary Computing* . C. Fogarty (Ed.), AISB Workshop, Leed , UK.
- Mitchell, M. (1996): *An Introduction to Genetic Algorithms*. The MIT Press.
- Shaw. K.J. and Fleming P.J., (1996): 'An initial study of practical multi-objective production scheduling using genetic algorithms. Proc. UKACC Int Conf on Control '96, Exeter, PP.448-453.
- Todd P.M. and Miller, G.F. (1991): On the Sympatric Origin of Species: Mercurial Mating in the Quicksilver Model. *Proceedings of the Fourth International Conference on Genetic Algorithms*. University of California, San Diego.
- Wright, S. (1932): The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proc. Sixth Int. Congr. Genetics*, 1, 356-366.
- Wright,S(1982):Character change,speciation, and the higher taxa.*Evolution*,36,427-443.